International Journal of Enterprise Information Systems
Volume 16 • Issue 4 • October-December 2020

Geographic Information System and Big Spatial Data:
A Review and Challenges

Raja Sher Afgun Usmani, Taylor's University, Malaysia
https://orcid.org/0000-0003-2027-1425
Ibrahim Abaker Targio Hashem, Taylor's University, Malaysia
Thulasyammal Ramiah Pillai, Taylor's University, Malaysia
Anum Saeed, Center for Advanced Studies in Engineering, Pakistan
Akibu Mahmoud Abdullahi, Taylor's University, Malaysia

ABSTRACT

Geographic information system (GIS) is designed to generate maps, manage spatial datasets, perform sophisticated "what if" spatial analyses, visualize multiple spatial datasets simultaneously, and solve location-based queries. The impact of big data is in every industry, including the GIS. The location-based big data also known as big spatial data has significant implications as it forces the industry to contemplate how to acquire and leverage spatial information. In this study, a comprehensive taxonomy is created to provide a better understanding of the uses of GIS and big spatial data. The taxonomy is made up of big data technologies, GIS data sources, tools, analytics, and applications. The authors look into the importance of big spatial data and its implications, review the data sources, and GIS analytics, applications, and GIS tools. Furthermore, in order to guide researchers interested in GIS, the challenges that require substantial research efforts are taken into account. Lastly, open issues in GIS that require further observation are summarized.

KEYWORDS

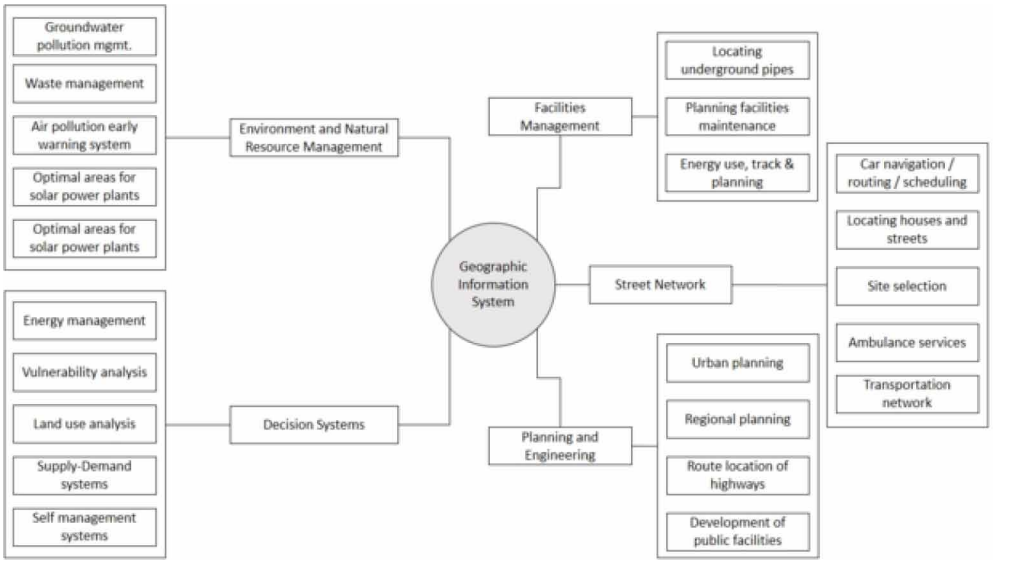
Analytics, Big Data, Big Spatial Data, Geographic Information Systems, GIS, GIS Tools, Prediction

INTRODUCTION

Geographic Information System (GIS) (Geography 2019) is to be a tool for preparing maps or generating presentation graphics. However, it is much more than that. GIS is an information system which aids the tool to gather data by collection, integration, management, analyzation, modeling and display for a reference to get accurate cartographic representation of an object in space (Chang 2016). Figure 1 provides an overview of GIS being applied around the world across many disciplines, professions and organizations. GIS has different characteristics that can separate it from other

DOI: 10.4018/IJEIS.2020100106

Figure 1. GIS Applications: Around the world



information systems. Primarily, GIS is designed for generating maps. These map-displays can be prepared using only a spreadsheet. It includes methods to digitize existing paper maps, collect coordinates of maps through surveying techniques or global positioning systems (GPS) and maps are generated using aerial photograph or satellite imagery maps.

Apart from generating maps, GIS is used to manage spatial datasets, i.e. GIS are spatial database management tools (Perumal *et al.* 2015). The data management process starts with outlining a link amongst the map and attribute data. The process of linking the attribute data with the spatial coordinates of the map is known as Geocoding. Geographic database is generated by creating the field in the attribute database for the longitude (location X-value) and latitude (location Y-value) of each address, this interdependently combines the map and attribute dataset.

The GIS provides an integrated environment to deal with the spatial criteria based queries, such as: "How many hotels are there within 2 km?" or "What is the air pollution index of Highbury, London, United Kingdom?". Additionally, GIS has the edge as it can answer the spatially referenced questions and concepts, such as "contained within" and "next to", that are usually not possible to answer with anecdotal information or database management systems. For instance, a user can perform attribute data-based queries by objects pointing, polygon definitions, or records selection from a particular distance of a dropped pin, whereas others do not. The GIS provides an integrated environment to deal with the spatial criteria based query. Spatially referenced questions such as "How many hotels are within 2 km?" or "what is the air pollution index of Highbury, London, United Kingdom". Additionally, GIS has the edge as it can answer the spatially referenced questions and concepts such as "contained within", "next to" that are usually not possible to answer with anecdotal information or database management systems. For instance, a user can perform attribute data based queries by objects pointing, by polygon definitions, or by records selection from a particular distance of a dropped pin whereas others do not.

Another capability of GIS is the visualization of spatial data. Using GIS, multiple datasets can be displayed on a map simultaneously because datasets are treated as different individual layers in GIS. These layers are explained as a table or a user-view. Layers can be stacked to show the relationship between them. A typical example of this can be a combination of a layer with school territories with the layer of enrollment information of schools in those territories. These datasets are stored as